

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

Shin TORIGOE, et al

Application. No.: *To Be Assigned*

Filed: *Concurrently*

FOR: ELECTRONIC DOCUMENT  
SIGNIFICANT UPDATING DETECTION  
APPARATUS, ELECTRONIC DOCUMENT  
SIGNIFICANT UPDATING DETECTION  
METHOD; ELECTRONIC DOCUMENT  
SIGNIFICANT UPDATING DETECTION  
PROGRAM, AND RECORDING MEDIUM ON  
WHICH ELECTRONIC DOCUMENT  
SIGNIFICANT UPDATING DETECTION  
PROGRAM IS RECORDED

Art Unit: *Unassigned*

Examiner: *Unassigned*

Atty. Docket No. 31759-190544

Customer No.



26694

PATENT TRADEMARK OFFICE

June 25, 2003

**SUBMISSION OF PRIORITY DOCUMENTS**

Commissioner for Patents  
P.O. Box 1450  
Alexandria, Virginia 22313-1450

Sir:

Submitted herewith are certified copies of Application No. JP2002-187859 filed on June 27, 2002 in Japan, and JP2003-55617 filed on March 3, 2003 in Japan, the priorities of which are claimed in the present application under the provisions of 35 U.S.C. 119. It is requested that the Examiner acknowledge receipt of the enclosed document in initial Office Action.

Respectfully submitted,

Norman N. Kunitz

Registration No. 20,586

VENABLE LLP

P.O. Box 34385

Washington, D.C. 20043-9998

Telephone: (202) 962-4800

Telefax: 202 (962-8300

NNK/elw  
#466237

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2003年 3月 3日

出 願 番 号

Application Number:

特願2003-055617

[ST.10/C]:

[JP2003-055617]

出 願 人

Applicant(s):

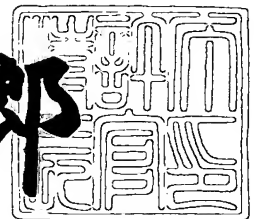
沖電気工業株式会社

Inventors: Shin TORIGOE, et al.  
ATTY REF: 31759-19054-1  
Customer No. 2003-11

2003年 4月 1日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

太田 信一郎



出証番号 出証特2003-3022449

【書類名】 特許願

【整理番号】 KN002565

【提出日】 平成15年 3月 3日

【あて先】 特許庁長官 太田 信一郎 殿

【国際特許分類】 G06F 15/00

【発明者】

【住所又は居所】 東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社  
社内

【氏名】 鳥越 真

【発明者】

【住所又は居所】 東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社  
社内

【氏名】 池野 篤司

【特許出願人】

【識別番号】 000000295

【氏名又は名称】 沖電気工業株式会社

【代表者】 篠塚 勝正

【代理人】

【識別番号】 100090620

【弁理士】

【氏名又は名称】 工藤 宣幸

【先の出願に基づく優先権主張】

【出願番号】 特願2002-187859

【出願日】 平成14年 6月27日

【手数料の表示】

【予納台帳番号】 013664

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9006358

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 電子文書有意更新検知装置、方法及びプログラム、並びに、電子文書有意更新検知プログラムを記録した記録媒体

【特許請求の範囲】

【請求項 1】 検知対象の電子文書及び比較対象の電子文書を取り込む入力手段と、

入力された検知対象の電子文書及び比較対象の電子文書の重要な箇所の差分を検知する有意更新検知手段とを有する

ことを特徴とする電子文書有意更新検知装置。

【請求項 2】 上記有意更新検知手段が、上記検知対象の電子文書及び上記比較対象の電子文書のそれぞれに対して、重要な箇所を抽出する前処理部と、上記前処理部により抽出された結果に対して、差分検知を行う差分検知部とを備えていることを特徴とする請求項 1 に記載の電子文書有意更新検知装置。

【請求項 3】 上記有意更新検知手段が、上記検知対象の電子文書及び上記比較対象の電子文書の間での差分検知を行う差分検知部と、抽出された差分に対して、有意な差分か否かを判断する価値判断部とを備えていることを特徴とする請求項 1 に記載の電子文書有意更新検知装置。

【請求項 4】 上記有意更新検知手段が、上記検知対象の電子文書及び上記比較対象の電子文書のそれぞれに対して、重要な箇所を抽出する前処理部と、上記前処理部により抽出された結果に対して、差分検知を行う差分検知部と、抽出された差分に対して、有意な差分か否かを判断する価値判断部とを備えていることを特徴とする請求項 1 に記載の電子文書有意更新検知装置。

【請求項 5】 上記有意更新検知手段の検知結果を、外部の情報処理装置に通知する出力手段をさらに備えていることを特徴とする請求項 1 ～ 4 のいずれかに記載の電子文書有意更新検知装置。

【請求項 6】 検知対象の電子文書及び比較対象の電子文書を取り込む入力工程と、

入力された検知対象の電子文書及び比較対象の電子文書の重要な箇所の差分を

検知する有意更新検知工程とを有する

ことを特徴とする電子文書有意更新検知方法。

【請求項 7】 上記有意更新検知工程が、上記検知対象の電子文書及び上記比較対象の電子文書のそれぞれに対して、重要な箇所を抽出する前処理と、この前処理により抽出された結果に対して、差分検知を行う差分検知処理とを有することを特徴とする請求項 6 に記載の電子文書有意更新検知方法。

【請求項 8】 上記有意更新検知工程が、上記検知対象の電子文書及び上記比較対象の電子文書の間での差分検知を行う差分検知処理と、抽出された差分に対して、有意な差分か否かを判断する価値判断処理とを有することを特徴とする請求項 6 に記載の電子文書有意更新検知方法。

【請求項 9】 上記有意更新検知工程が、上記検知対象の電子文書及び上記比較対象の電子文書のそれぞれに対して、重要な箇所を抽出する前処理と、この前処理により抽出された結果に対して、差分検知を行う差分検知処理と、抽出された差分に対して、有意な差分か否かを判断する価値判断処理とを含むことを特徴とする請求項 6 に記載の電子文書有意更新検知方法。

【請求項 10】 上記有意更新検知工程の検知結果を、外部の情報処理装置に通知する出力処理をさらに有することを特徴とする請求項 6 ～ 9 のいずれかに記載の電子文書有意更新検知方法。

【請求項 11】 請求項 6 ～ 10 のいずれかに記載の電子文書有意更新検知方法の各工程をコンピュータが処理し得るコードで記述したことを特徴とする電子文書有意更新検知プログラム。

【請求項 12】 請求項 11 の電子文書有意更新検知プログラムを記録していることを特徴とする記録媒体。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、電子文書有意更新検知装置、方法及びプログラム、並びに、電子文書有意更新検知プログラムを記録した記録媒体に関し、例えば、Web ページやテキスト等の電子文書の更新を監視し、更新の旨をユーザへ通知するシステムに

適用し得るものである。

【 0 0 0 2 】

【従来の技術】

従来、同一のURLに係るWebページも適宜更新される。このようなWebページの更新を検出する方式として、特許文献1に開示されたものがあり、対象とするWebページのチェックサムを比較し、変化があればWebページが更新されたとみなしている。

【 0 0 0 3 】

【特許文献1】 特開2000-35913号公報

【 0 0 0 4 】

【発明が解決しようとする課題】

しかしながら、上述の方式では、文章の微小な修正（例えば、誤字、脱字、訂正等）や、本文とは関連のない部位（例えば、広告欄や他の小さな見出し等）の更新までも、Webページの更新として検知してしまうため、有意な更新を期待する多くの利用者にとって、不要な結果が得られることになる。

【 0 0 0 5 】

そのため、電子文書の更新と呼ぶに相応しいようなレベルの更新を検知できる電子文書有意更新検知装置等が望まれている。

【 0 0 0 6 】

【課題を解決するための手段】

第1の本発明の電子文書有意更新検知装置は、検知対象の電子文書及び比較対象の電子文書を取り込む入力手段と、入力された検知対象の電子文書及び比較対象の電子文書の重要な箇所の差分を検知する有意更新検知手段とを有することを特徴とする。

【 0 0 0 7 】

第2の本発明の電子文書有意更新検知方法は、検知対象の電子文書及び比較対象の電子文書を取り込む入力工程と、入力された検知対象の電子文書及び比較対象の電子文書の重要な箇所の差分を検知する有意更新検知工程とを有することを特徴とする。

【 0 0 0 8 】

第 3 の本発明の電子文書有意更新検知プログラムは、第 2 の本発明の電子文書有意更新検知方法の各工程をコンピュータが処理し得るコードで記述したことを特徴とする。

【 0 0 0 9 】

第 4 の本発明の記録媒体は、第 3 の本発明の電子文書有意更新検知プログラムを記録していることを特徴とする。

【 0 0 1 0 】

【発明の実施の形態】

(A) 第 1 の実施形態

以下、本発明による電子文書有意更新検知装置、方法及びプログラム、並びに、電子文書有意更新検知プログラムを記録した記録媒体の第 1 の実施形態を図面を参照しながら詳述する。

【 0 0 1 1 】

(A-1) 第 1 の実施形態の構成

図 1 は、第 1 の実施形態の電子文書有意更新検知装置の機能的構成を示すブロック図である。

【 0 0 1 2 】

例えば、第 1 の実施形態の電子文書有意更新検知装置は、通信機能を有するユーザのパソコン等の情報処理装置又はプロバイダサーバ等の上で実現されるが、機能的には、図 1 で表すことができる。例えば、パソコン等の情報処理装置又はプロバイダサーバ等に対し、CD-ROM やフレキシブルディスク等の記録媒体上に記録されている電子文書有意更新検知プログラムをインストールすることにより、第 1 の実施形態の電子文書有意更新検知装置が構築される。なお、実際上、1 つのシステム上で構築しても構わないし、ネットワークで結ばれたサーバ間で協調動作するように構築しても構わない。

【 0 0 1 3 】

第 1 の実施形態の電子文書有意更新検知装置は、入力部 1、有意更新検知部 2 及び出力部 5 を有し、有意更新検知部 2 は、前処理部 3 と差分抽出部 4 とを有す



る。

【 0 0 1 4 】

入力部 1 は、インターネットやイントラネット等のネットワークから、又は、CD-ROM等の記録媒体から、Web ページやテキスト等の電子文書を取得して入力データとする。

【 0 0 1 5 】

入力部 1 は、有意な更新の検知対象である電子文書及び比較対象の電子文書の 2 文書を、バージョンを指定して取り出すことが可能ならば、同時に取り出しても良く、また、URL を指定して過去に取り出し記憶していたものを比較対象の電子文書とし、同一の URL で今回取り出したものを有意な更新の検知対象の電子文書として取り出しても良く、さらには、異なる時期の過去に取り出して記憶していた新旧の 2 文書を検知対象である電子文書及び比較対象の電子文書として入力しても良い。

【 0 0 1 6 】

有意更新検知部 2 は、比較対象の電子文書に対する検知対象である電子文書の有意な更新部分を検知するものである。有意更新検知部 2 において、前処理部 3 は、各電子文書のそれぞれに対して、重要な箇所を抽出するものであり、差分抽出部 4 は、前処理部 3 によって抽出された重要な箇所同士におけるテキスト列の差分を抽出するものである。

【 0 0 1 7 】

ここで、電子文書の重要な箇所とは、例えば、電子文書の本文、又は、本文やタイトル内の主要文章（その要約を含む）を言い、これらとは関連のない部位（例えば、広告欄や他の小さな見出し等）は非重要な箇所とする。

【 0 0 1 8 】

前処理部 3 による電子文書の重要な箇所の抽出方法としては、既存の方法を適用することができる。また、重要な箇所を、自動判定しても良く、重要な箇所をユーザが特定しても良い。

【 0 0 1 9 】

例えば、Web ページは、HTML や XML 等で記述されており、複数のフレ

ームによって1画像が形成されるが、各フレーム部分を規定するタグ識別子（例えば「MAIN」）や、各フレーム部分の面積の大小や、各フレーム部分における文字数や、各フレームの配置位置や、所定のキーワードを含むか否か等によって、重要な箇所（フレーム部分）を判別することができる。

#### 【0020】

差分抽出部4におけるテキスト列の差分抽出方法としても、既存の方法を適用することができる。

#### 【0021】

出力部5は、Webページ等の電子文書に有意な更新であった場合に、その旨を表示装置に表示したり、ユーザに電子メールで更新内容を通知したりするものである。出力内容には、更新前後の内容を含んでいても良く、また、更新された部分を特定した更新後の内容であっても良く、また、その出力形式も任意で構わない。

#### 【0022】

##### （A-2）第1の実施形態の動作

以下、第1の実施形態の詳細な処理について、架空の更新前後のWebページを例として取り上げて説明する。なお、図2は、更新前のWebページを示し、図3は、更新後のWebページを示している。また、上述した図1は、機能ブロック図であったが、処理の流れを示すフローチャートと見ることもできる。

#### 【0023】

符号11は、更新前のWebページのブラウザでの表示例を表しており、符号16は、更新後のWebページのブラウザでの表示例を表している。更新後のWebページ16には、更新箇所を明示するための便宜上、更新された個所に下線を付与しているが、Webページ自体にはこの下線は付与されていない。

#### 【0024】

更新前後のWebページ11及び16は、4個のフレーム12～15（図2参照）でなっており、それぞれ、ヘッダ、メニュー、記事、フッタとする。

#### 【0025】

入力部1は、図2及び図3に示す更新前後のWebページ11及び16を取り

込んで有意更新検知部 2 に与える。

【0026】

有意更新検知部 2 は、前処理部 3 と差分検知部 4 を含んでおり、前処理部 3 において、対象となる文書に対して重要個所の抽出を行い、その結果を差分検知部 4 で比較する。

【0027】

前処理部 3 による重要個所の抽出方法は、例えば、ユーザによるフレームの事前指定、要約（重要文抽出）等の種々の方法があるが、以下では、ユーザによるフレームの事前指定方法を用いた例と、要約（重要文抽出）を抽出する例とを説明する。

【0028】

ユーザによるフレームの事前指定は、図 4 に示すような注目個所テーブルを用いて、ユーザが監視して欲しい Web ページの URL と更新を検知して欲しい個所（フレーム）を指定しておく。前処理部 3 では、この情報に基づき、対象となる Web ページ内の特定フレームを抽出してその特定フレームのみを差分抽出部 4 へ送る。このときの処理イメージを図 5 に示す。フレーム群 17 は、図 4 で指定されなかったフレームの一群を示しており、フレーム 18 は図 4 で指定して抽出されたフレームを示している。図 5 は更新後の Web ページの抽出イメージを示しているが、図示は省略するが、更新前の Web ページに対しても同様な抽出が行われる。

【0029】

差分抽出部 4 では、更新前後の Web ページにおけるフレーム 18 同士の差分のみを抽出する。図 5 に示すフレーム 18 の下線部分は、更新後の Web ページにおける差分抽出部 4 によって抽出された差分部分を示している。

【0030】

一方、要約抽出（重要文抽出）は、文書内の文字列から重要と思われる文を抽出する手法であり、例えば、特開平 11-272686 号公報に開示されるものを適用できる。前処理部 3 では重要と思われる文字列（文）を抽出して差分抽出部 4 へ送る。

## 【 0 0 3 1 】

このときの処理イメージを図 6 に示す。図 6 において、符号 1 9、2 0 を付与したものはそれぞれ、前処理部 3 による更新前と更新後の W e b ページの要約抽出結果を示している。なお、図 6 の処理イメージ図 1 9 及び 2 0 において、それぞれ重要でないと判断された文字列は二重線で消去して示しているが、これは、理解の容易化のためであり、これらの文字列は、重要でないため抽出されず、差分抽出部 4 には引き渡されない。

## 【 0 0 3 2 】

図 6 において、符号 2 1 を付与したものは、差分抽出部 4 による差分抽出結果を示している。差分抽出部 4 は、重要文として抽出された二重線で消去されていない文同士を比較照合し、符号 2 1 を付した文で下線が付与された部分を差分であると抽出する。なお、図 6 の処理イメージ図 2 1 において、差分抽出部分に下線を付与しているが、これは、理解の容易化のためであり、文字列に対する下線付与動作を、差分抽出部 4 は必ずしも実行していない。

## 【 0 0 3 3 】

前処理部 3 の別の手法（追加手法）としては、キーワード抽出による微小な修正等の除去も挙げることができる。キーワード抽出は、例えば、キーワードを「異なる文字コードに囲まれた漢字及びカナの連続文字」として定義することにより、上述した図 2 及び図 3 に示す更新前後の W e b ページに対するキーワード抽出結果は、図 7 に示すようになる。更新前後の W e b ページのフレーム 1 3 及び 1 5 の変更箇所（「サイトマップ」や「e - m a i l」）は、上述の定義ではキーワードとはなり得ないため抽出されない。図 7 に示すようなキーワード抽出結果を、差分抽出部 4 で比較することで、更新されたかどうかを判別することができる。なお、キーワード抽出だけを用いた場合において、図 2 及び図 3 のフレーム 1 4 内の 1 月 1 日の記事における「します」を「しました」にする変更のみがあったときには、変更前と変更後のキーワードの違いは生じず、微小な修正であり、有意な更新ではないと判別される。

## 【 0 0 3 4 】

出力部 5 では、差分抽出部 4 の結果に基づき、対象とする W e b ページに有意

な更新があることを出力する。例えば、対象とするW e b ページに有意な更新があったことをユーザに通知する。

【 0 0 3 5 】

ユーザへの通知は、表示デバイスへの表示やメールによる通知などで行うことができ、通知内容は、対象となったW e b ページのU R L や、変更を検知したフレームの情報でもよく、具体的な変更内容を含めても構わない。また、ユーザへの通知は、ユーザが該当するW e b ページを取り出そうとしたタイミングで行っても良い。

【 0 0 3 6 】

なお、更新前のW e b ページの情報を予め格納しておくバッファや、任意のタイミングで対象W e b ページを取得するタイマー類の存在等は、容易に理解できるので、その説明は省略している。また、格納しておく更新前のW e b ページの情報は、W e b ページの生のデータでも良く、前処理部 3 の処理後のデータでも良い。

【 0 0 3 7 】

(A-3) 第 1 の実施形態の効果

以上のように、第 1 の実施形態によれば、前処理部 3 において対象となる更新前後の電子文書に対して重要個所の抽出を行い、差分抽出部 4 では、重要個所の変化を有意な更新として検知することができる。これにより、出力部 5 において、有意な更新があった旨をユーザに伝えることが可能となる。

【 0 0 3 8 】

なお、前処理部 3 において、キーワード抽出を利用することにより、差分抽出部 4 では微小な修正は検知対象外とでき、真に有意な更新のみを検知することができる。

【 0 0 3 9 】

(B) 第 2 の実施形態

次に、本発明による電子文書有意更新検知装置、方法及びプログラム、並びに、電子文書有意更新検知プログラムを記録した記録媒体の第 2 の実施形態を図面を参照しながら詳述する。

## 【 0 0 4 0 】

## ( B - 1 ) 第 2 の実施形態の構成

図 8 は、第 2 の実施形態の電子文書有意更新検知装置の機能的構成を示すブロック図である。

## 【 0 0 4 1 】

例えば、第 2 の実施形態の電子文書有意更新検知装置も、通信機能を有するユーザのパソコン等の情報処理装置又はプロバイダサーバ等の上で実現されるが、機能的には、図 8 で表すことができる。記録媒体上の電子文書有意更新検知プログラムをインストールして、第 2 の実施形態の電子文書有意更新検知装置を構築しても良い。なお、實際上、1つのシステム上で構築しても構わないし、ネットワークで結ばれたサーバ間で協調動作するように構築しても構わない。

## 【 0 0 4 2 】

第 2 の実施形態の電子文書有意更新検知装置も、大きくは、入力部 1、有意更新検知部 6 及び出力部 5 を有するが、有意更新検知部 6 の内部構成が第 1 の実施形態のものと異なっており、入力部 1 及び出力部 5 は、第 1 の実施形態のものと同様である。

## 【 0 0 4 3 】

第 2 の実施形態の有意更新検知部 6 も、Web ページ等の電子文書の有意な更新を検知するものであが、第 2 の実施形態の有意更新検知部 6 は、差分抽出部 4 及び価値判断部 7 を有している。

## 【 0 0 4 4 】

差分抽出部 4 は、第 1 の実施形態と同様な方法によって差分を抽出するものであるが、第 2 の実施形態の場合、差分抽出対象は、電子文書全体になっている点が、第 1 の実施形態と異なっている。

## 【 0 0 4 5 】

価値判断部 7 は、差分抽出部 4 が抽出した差分が有意なものであるかを判断し、有意な差分だけを抽出するものである。価値判断部 7 は、例えば、差分量（例えば差分文字数）の閾値との比較処理や、形態素解析等の自然言語処理による属性判断等を用いて有意な差分を判別する。

【 0 0 4 6 】

( B - 2 ) 第 2 の実施形態の動作

第 2 の実施形態の詳細な処理についても、上述した図 2 及び図 3 で示す架空の更新前後の W e b ページを例として取り上げて説明する。

【 0 0 4 7 】

上述のように有意更新検知部 6 は、差分抽出部 4 と価値判断部 7 を含んでおり、差分抽出部 4 によって文書全体の差分の抽出を行い、その結果の有意性を価値判断部 7 によって判断する。

【 0 0 4 8 】

第 2 の実施形態の場合、差分抽出対象が電子文書全体になっている点が第 1 の実施形態とは異なっているが、差分抽出部 4 による差分抽出方法自体は、第 1 の実施形態と同様であるので、その説明は省略し、以下では、価値判断部 7 による差分価値判断処理を詳述する。なお、図 9 における符号 2 2 を付したものは、図 2 及び図 3 に示す更新前後の W e b ページに対し、第 2 の差分抽出部 4 が抽出した差分結果を示すものである。

【 0 0 4 9 】

価値判断部 7 による差分価値判断処理を、以下では、差分量の閾値との比較処理を用いた差分価値判断処理と、形態素解析等の自然言語処理による属性判断を用いた差分価値判断処理とについて説明する。

【 0 0 5 0 】

差分量の閾値との比較処理を用いた差分価値判断処理では、例えば、個々の差分の文字列長（例えば文字数又は全角に置き直した文字数等）がある閾値を超えた場合に価値がある（有意な差分である）と判断するものである。

【 0 0 5 1 】

仮に、差分の文字列が 1 0 文字以上で有効（有意）とした場合（閾値が 1 0 文字）であれば、図 9 の差分抽出結果における差分「サイトマップ」、「した」、「e - m a i l .」は有意でないと判断され、一方、差分「2 月…開催します。」は有意であると判断される。その結果、価値判断部 7 による判断結果は、図 9 に符号 2 3 で示した部分の二重線が付与されていない文字列になる。言い換える

と、閾値未満の文字列を削除（二重線部分参照）することで、価値判断部 7 は、まとまりのある文に価値があると判別している。

#### 【 0 0 5 2 】

形態素解析等の自然言語処理による属性判断を用いた差分価値判断処理では、差分抽出部 4 から渡された図 9 に示すような差分 2 2 をいくつかのパーツに分類し、各パーツの属性に基づいて、価値（有意な差分）を判別していく。例えば、文を構成しないパーツ（例えば助詞や単独の名詞等）を不要パーツと定義して判別する。この場合の判別結果も、図 9 の符号 2 3 を付与した内容で表され、不要パーツ（二重線参照）を削除することにより、まとまりのある文に価値があると判別している。なお、月日は、その後のスペースを挟んで文章に繋がっているときには、文章の一部と認識するようにしている。

#### 【 0 0 5 3 】

価値判断部 7 によって価値がある（有意な箇所）と判断された文字列は出力部 5 に渡され、第 1 の実施形態と同様にして出力される。

#### 【 0 0 5 4 】

なお、第 2 の実施形態の説明においても、更新前の W e b ページの情報を予め格納しておくバッファや、任意のタイミングで対象 W e b ページを取得するタイマー類の存在等は、容易に理解できるので、その説明は省略している。

#### 【 0 0 5 5 】

##### （B－3）第 2 の実施形態の効果

以上のように、第 2 の実施形態によれば、価値判断部 7 において対象となる文書の差分文字列に対して価値判断を行うことで、文書の小さな修正等を更新情報から除外することができる。これにより、有意更新検知部 6 では対象となる文書の更新内容のうち有意な情報のみを検知し、出力部 5 においてその更新内容をユーザ等に出力することができる。

#### 【 0 0 5 6 】

##### （C）他の実施形態

第 1 の実施形態及び第 2 の実施形態は、インターネット、イントラネットの W e b ページやテキスト文書を監視するシステムに利用することができる。このよ



うにした場合には、システム側では多くのユーザからの個別アクセスによるトラフィックを減らすことができると共に、ユーザ側ではサイトの巡回を行う時間と労力を削減することが可能となる。

【 0 0 5 7 】

第 1 及び第 2 の実施形態共に、有意な更新があったかどうかを検知してその旨を出力するようにしても良く、有意と判断した情報そのものを出力するようにしても良い。

【 0 0 5 8 】

第 1 の実施形態の技術思想と第 2 の実施形態の技術思想は、個別にシステムに実装しても良く、同時にシステムに実装しても良い。

【 0 0 5 9 】

また、第 1 の実施形態の前処理部 3 で用いた処理を、第 2 の実施形態の価値判別部 7 の処理にアレンジしても良く、逆に、第 2 の実施形態の価値判別部 7 で用いた処理を、第 1 の実施形態の前処理部 3 の処理にアレンジしても良い。これらの工夫で、処理の強化や、サイト毎のきめ細かな処理に対応することが可能となる。

【 0 0 6 0 】

さらに、上記各実施形態は、更新後の電子文書での更新情報を出力するものを意図しているが、更新前の電子文書での更新情報を出力するものであっても良く、両者の更新情報を出力するものであっても良い。

【 0 0 6 1 】

さらにまた、有意差分を抽出するための 2 電子文書は、任意の時期のものであって良く、一方が最新のものに限定されるものではない。

【 0 0 6 2 】

なお、上記では、差分が抽出できた例を説明したが、差分が存在しない場合には、その旨を出力するようにしても良い。また、出力が、ユーザに通知する態様の場合には、ユーザに通知しないようにしても良い。さらに、差分が、一方の電子文書の全体又は所定フレームの全体であれば、両文書が比較照合をする関係ないことを出力するようにしても良い。

【 0 0 6 3 】

【発明の効果】

以上のように、本発明によれば、電子文書の更新と呼ぶに相応しいようなレベルの更新を検知することができる。

【図面の簡単な説明】

【図 1】

第 1 の実施形態の電子文書有意更新検知装置の機能的構成を示すブロック図である。

【図 2】

更新前の W e b ページ例を示す説明図である。

【図 3】

図 2 の W e b ページに対応した更新後の W e b ページ例を示す説明図である。

【図 4】

第 1 の実施形態のフレーム事前指定に利用する注目箇所テーブルを示す説明図である。

【図 5】

第 1 の実施形態の W e b ページでの注目フレームの説明図である。

【図 6】

第 1 の実施形態の要約（重要文）の抽出方法の説明図である。

【図 7】

第 1 の実施形態のキーワード抽出という前処理で得られたキーワード例の説明図である。

【図 8】

第 2 の実施形態の電子文書有意更新検知装置の機能的構成を示すブロック図である。

【図 9】

第 2 の実施形態の動作の説明図である。

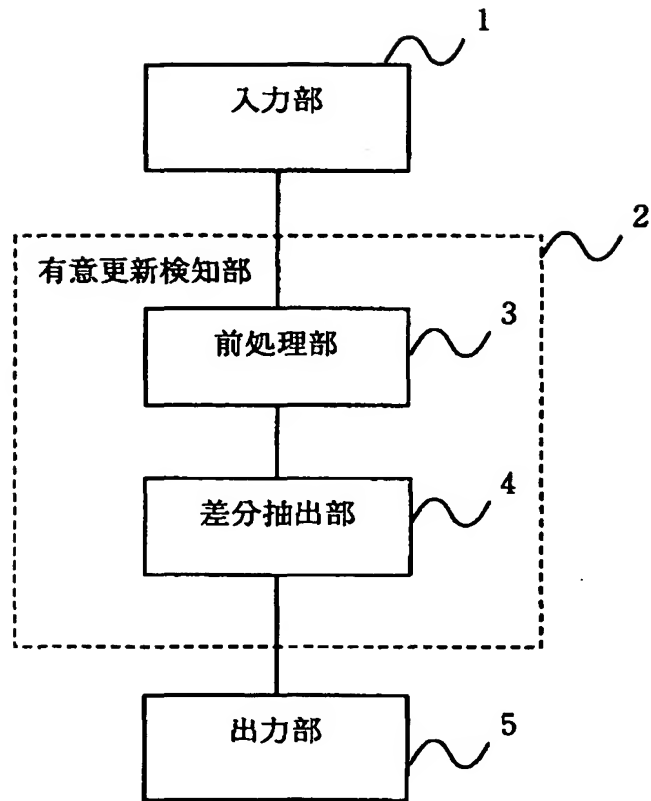
【符号の説明】

1 …入力部、 2、 6 …有意更新検知部、 3 …前処理部、 4 …差分抽出部、 5 …

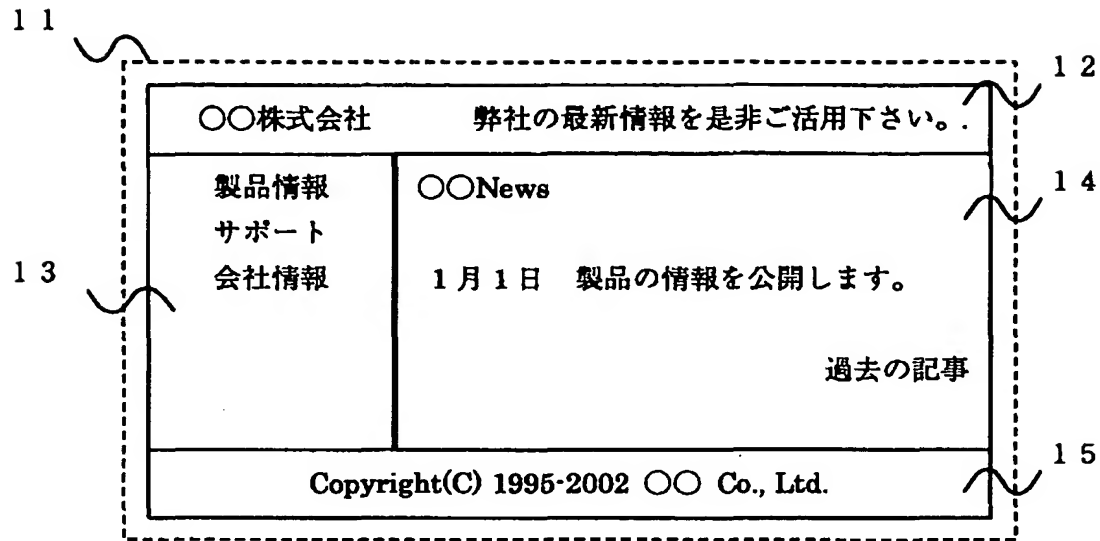
出力部、7…価値判断部。

【書類名】 図面

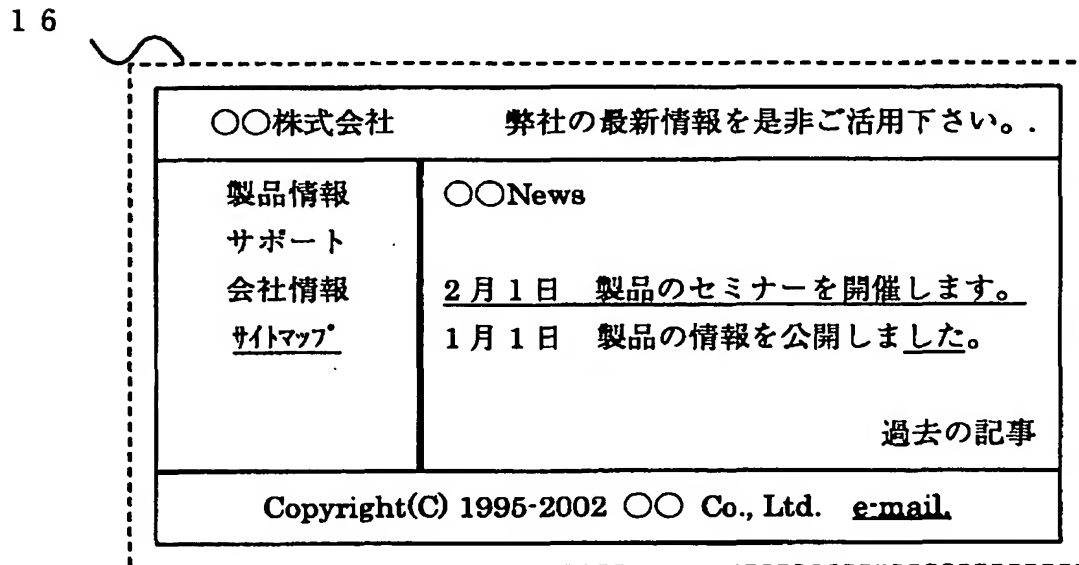
【図 1】



【図 2】



【図 3】



【図 4】

URL	frame
http://www.〇〇.jp/	MAIN
http://...	...

【図 5】

〇〇株式会社

製品情報

サポート

会社情報

お付まつ

弊社の最新情報を是非ご利用下さい。

〇〇News

2月1日 製品のセミナーを開催します。

1月1日 製品の情報を公開しました。

過去の記事

17

18

Copyright(©) 1995-2002 〇〇 Co., Ltd. e-mail

【図 6】

〇〇株式会社 弊社の最新情報を是非ご利用下さい。.	
製品情報 <del>サポート</del> 会社情報	〇〇News 1月1日 製品の情報を公開します。 <del>過去の記事</del>
Copyright(C) 1995-2002 〇〇 Co., Ltd.	

19

〇〇株式会社 弊社の最新情報を是非ご利用下さい。.	
製品情報 <del>サポート</del> 会社情報 <del>#11777</del>	〇〇News 2月1日 製品のセミナーを開催します。 1月1日 製品の情報を公開しました。 <del>過去の記事</del>
Copyright(C) 1995-2002 〇〇 Co., Ltd. e-mail	

20

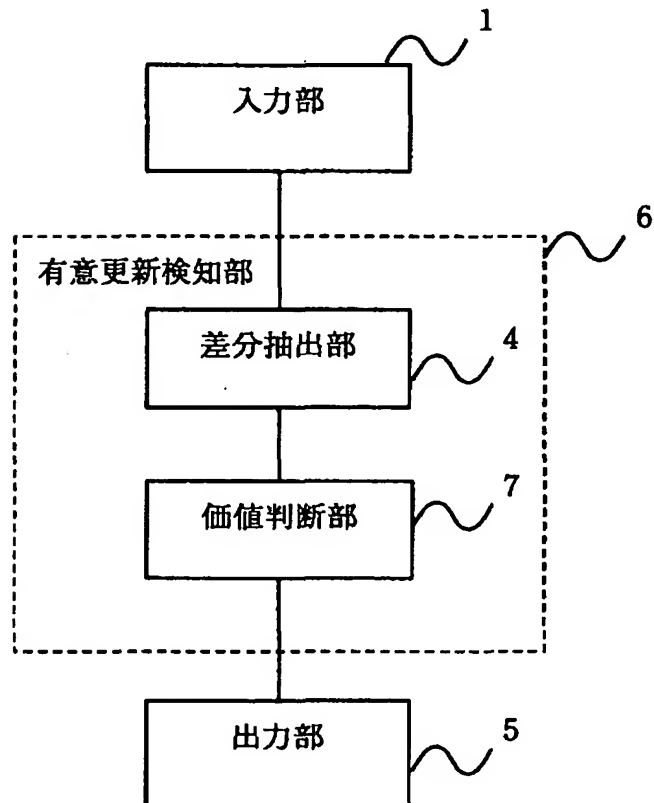
弊社の最新情報を是非ご利用下さい。.	
<u>2月1日 製品のセミナーを開催します。</u> <u>1月1日 製品の情報を公開しました。</u>	

21

【図 7】

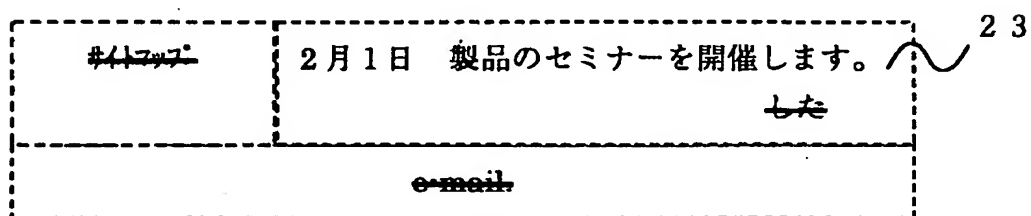
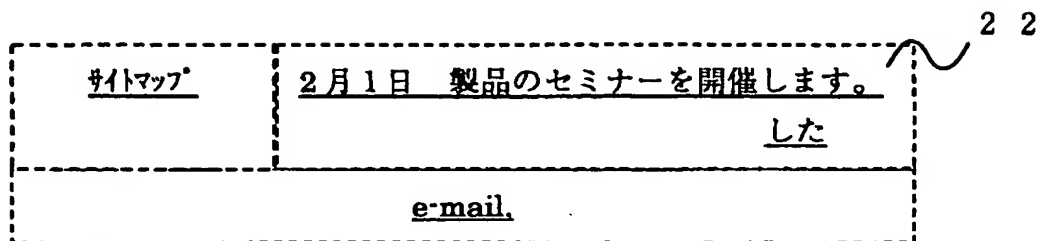
更新前のキーワード	更新後のキーワード
製品	製品
情報	情報
公開	公開
弊社	弊社
最新情報	最新情報
活用	活用
	セミナー
	開催

【図 8】





【図 9】



【書類名】 要約書

【要約】

【課題】 電子文書の更新と呼ぶに相応しいようなレベルの更新箇所を検知する

。

【解決手段】 本発明では、検知対象の電子文書及び比較対象の電子文書を取り込み、検知対象の電子文書及び比較対象の電子文書の重要な箇所の差分を検知する。重要な箇所の差分は、（１）各電子文書の重要な箇所を抽出した後、差分検知を行って得たり、（２）両電子文書全体の差分を検知した後、有意な差分か否かを判断したり、（３）各電子文書の重要な箇所を抽出した後、差分検知を行い、さらに、有意な差分か否かを判断したりすることで得る。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [ 0 0 0 0 0 0 2 9 5 ]

1. 変更年月日 1 9 9 0 年 8 月 2 2 日  
[変更理由] 新規登録  
住 所 東京都港区虎ノ門1丁目7番12号  
氏 名 沖電気工業株式会社

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2002年 6月27日

出 願 番 号

Application Number:

特願2002-187859

[ ST.10/C ]:

[ JP2002-187859 ]

出 願 人

Applicant(s):

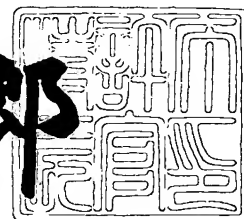
沖電気工業株式会社

Inventors: Shin TORIGOE et al.  
ATTY DKT: 31759-190544  
Customer No. 26694

2003年 4月 1日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

太田 信一郎



出証番号 出証特2003-3022444

【書類名】 特許願

【整理番号】 KN002527

【提出日】 平成14年 6月27日

【あて先】 特許庁長官 及川 耕造 殿

【国際特許分類】 G06F 15/00

【発明者】

【住所又は居所】 東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社  
社内

【氏名】 鳥越 真

【発明者】

【住所又は居所】 東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社  
社内

【氏名】 池野 篤司

【特許出願人】

【識別番号】 000000295

【氏名又は名称】 沖電気工業株式会社

【代表者】 篠塚 勝正

【代理人】

【識別番号】 100090620

【弁理士】

【氏名又は名称】 工藤 宣幸

【手数料の表示】

【予納台帳番号】 013664

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9006358

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 電子文書有意更新検知装置および電子文書有意更新通知装置

【特許請求の範囲】

【請求項 1】 複数の電子文書のそれぞれに対して、重要な箇所を抽出する 1 つ以上の前処理手段と、

前処理により抽出された結果に対して差分検知を行う差分検知手段と

を備えていることを特徴とする電子文書有意更新検知装置。

【請求項 2】 複数の電子文書それぞれに対して差分検知を行う差分検知手段と、

抽出された差分に対して有意かどうかの価値を判断する 1 つ以上の価値判断手段と

を備えることを特徴とする電子文書有意更新検知装置。

【請求項 3】 複数の電子文書のそれぞれに対して重要な箇所を抽出する 1 つ以上の前処理手段と、

それぞれの電子文書に対して差分検知を行う差分検知手段と、

抽出された差分に対して有意かどうかの価値を判断する 1 つ以上の価値判断手段と

を備えることを特徴とする電子文書有意更新検知装置。

【請求項 4】 上記前処理手段が、電子文書の主要箇所を取り出すことができる主要箇所抽出手段を備えることを特徴とする請求項 1 又は 3 に記載の電子文書有意更新検知装置。

【請求項 5】 上記価値判断手段が、電子文書の主要箇所を取り出すことができる主要箇所抽出手段を備えることを特徴とする請求項 2 又は 3 に記載の電子文書有意更新検知装置。

【請求項 6】 上記前処理手段が、電子文書のキーワードを取り出すキーワード抽出手段を備えることを特徴とする請求項 1 又は 3 に記載の電子文書有意更新検知装置。

【請求項 7】 上記価値判断手段が、電子文書のキーワードを取り出すキー

ワード抽出手段を備えることを特徴とする請求項 2 又は 3 に記載の電子文書有意更新検知装置。

【請求項 8】 上記前処理手段が、電子文書の要約を生成することができる要約生成手段を備えることを特徴とする請求項 1 又は 3 に記載の電子文書有意更新検知装置。

【請求項 9】 上記価値判断手段が、電子文書の要約を生成することができる要約生成手段を備えることを特徴とする請求項 2 又は 3 に記載の電子文書有意更新検知装置。

【請求項 10】 上記前処理手段が、形態素解析などの自然言語処理を行う手段を備えることを特徴とする請求項 1 又は 3 に記載の電子文書有意更新検知装置。

【請求項 11】 上記価値判断手段が、形態素解析などの自然言語処理を行う手段を備えることを特徴とする請求項 2 又は 3 に記載の電子文書有意更新検知装置。

【請求項 12】 上記価値判断手段が、差分量が一定の閾値を超えているか否かで価値を判断する差分量比較手段を備えることを特徴とする請求項 2 又は 3 に記載の電子文書有意更新検知装置。

【請求項 13】 電子文書の更新を検知して通知する電子文書有意更新通知装置において、

請求項 1 ～ 12 のいずれかに記載の電子文書有意更新検知装置を備え、その検知結果を通知することを特徴とする電子文書有意更新通知装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、電子文書有意更新検知装置および電子文書有意更新通知装置に関し、例えば、Web ページやテキスト等の電子文書の更新を監視し、更新の旨をユーザへ通知するシステムに適用し得るものである。

【0002】

【従来の技術】



従来、この種の方式として、特開 2 0 0 0 - 3 5 9 1 3 号公報に開示されたものがあり、対象とする Web ページのチェックサムを比較し、変化があれば Web ページが更新されたとみなしている。

【 0 0 0 3 】

【発明が解決しようとする課題】

しかしながら、上記方式では、文章の微小な修正（例えば、誤字、脱字、訂正等）や、本文とは関連のない部位（例えば、広告欄や他の小さな見出し等）までも「更新」として検知してしまうため、有意な更新を期待する多くの利用者にとって、不要な通知が届くことになる。

【 0 0 0 4 】

そのため、電子文書の更新と呼ぶに相応しいようなレベルの更新を検知したり通知したりできる電子文書有意更新検知装置や電子文書有意更新通知装置が望まれている。

【 0 0 0 5 】

【課題を解決するための手段】

第 1 の本発明の電子文書有意更新検知装置は、複数の電子文書のそれぞれに対して、重要な箇所を抽出する 1 つ以上の前処理手段と、前処理により抽出された結果に対して差分検知を行う差分検知手段とを備えていることを特徴とする。

【 0 0 0 6 】

第 2 の本発明の電子文書有意更新検知装置は、複数の電子文書それぞれに対して差分検知を行う差分検知手段と、抽出された差分に対して有意かどうかの価値を判断する 1 つ以上の価値判断手段とを備えることを特徴とする。

【 0 0 0 7 】

第 3 の本発明の電子文書有意更新検知装置は、複数の電子文書のそれぞれに対して重要な箇所を抽出する 1 つ以上の前処理手段と、それぞれの電子文書に対して差分検知を行う差分検知手段と、抽出された差分に対して有意かどうかの価値を判断する 1 つ以上の価値判断手段とを備えることを特徴とする。

【 0 0 0 8 】

第 4 の本発明は、電子文書の更新を検知して通知する電子文書有意更新通知装

置において、第 1、第 2 又は第 3 の本発明の電子文書有意更新検知装置を備え、その検知結果を通知することを特徴とする。

【 0 0 0 9 】

【発明の実施の形態】

(A) 第 1 の実施形態

以下、本発明による電子文書有意更新検知装置および電子文書有意更新通知装置の第 1 の実施形態を図面を参照しながら詳述する。

【 0 0 1 0 】

(A-1) 第 1 の実施形態の構成

図 1 は、第 1 の実施形態の電子文書有意更新通知装置の機能的構成を示すブロック図である。例えば、第 1 の実施形態の電子文書有意更新通知装置は、通信機能を有するパソコン等の情報処理装置又はプロバイダーサーバ等の上で実現されるが、機能的には、図 1 で表すことができる。なお、實際上、1 つのシステム上で構築してもかまわないし、ネットワークで結ばれたサーバ間で協調動作するように構築してもかまわない。

【 0 0 1 1 】

第 1 の実施形態の電子文書有意更新通知装置は、入力部 1、有意更新検知部（第 1 の実施形態の電子文書有意更新検知装置）2 および出力部 5 を有し、有意更新検知部 2 は、前処理部 3 と差分抽出部 4 とを有する。

【 0 0 1 2 】

入力部 1 は、インターネットやイントラネット等から Web ページやテキスト等を取得して入力データとする。

【 0 0 1 3 】

有意更新検知部 2 は、Web ページ等の有意な更新を検知し、その内部では、前処理部 3 による文書の前処理と、差分抽出部 4 による前処理処理済み文書の差分抽出を行うものである。

【 0 0 1 4 】

出力部 5 では、Web ページ等に有意な更新であった場合に、その旨を表示装置に表示したり、ユーザにメールで更新内容を通知したりするものである。

【 0 0 1 5 】

(A-2) 第 1 の実施形態の動作

第 1 の実施形態の詳細な処理について、架空の W e b ページ (図 2 と図 3) を例として取り上げて説明する。なお、図 2 は、更新前の W e b ページを示し、図 3 は、更新後の W e b ページを示している。

【 0 0 1 6 】

1 1 は更新前の W e b ページのブラウザでの表示例であり、1 6 は更新後の表示例である。1 6 では、便宜上、更新された個所に下線を付与している。

【 0 0 1 7 】

1 2 ~ 1 5 は W e b ページでのフレームを表しており、それぞれ、ヘッダ、メニュー、記事、フッタとする。

【 0 0 1 8 】

有意更新検知部 2 は、前処理部 3 と差分検知部 4 を含んでおり、前処理部 3 において、対象となる文書に対して主要個所の抽出を行い、その結果を差分検知部 4 で比較する。前処理部 3 の主要個所の抽出方法は、例えば、ユーザによるフレームの事前指定、要約 (重要文抽出)、等がある。以下、これらの手法を用いた例を説明する。

【 0 0 1 9 】

ユーザによるフレームの事前指定は、図 4 に示すような注目個所テーブルを用いて、ユーザが監視して欲しい W e b ページの U R L と更新を検知して欲しい個所 (フレーム) を指定しておく。前処理部 3 では、この情報に基づき、対象となる W e b ページ内の特定フレームのみを差分検出部 4 へ送る。このときの処理イメージを図 5 へ示す。1 7 は図 4 で指定されなかったフレーム群、1 8 は図 4 で指定したフレームを表す。差分検出部 4 では、1 8 のフレームのみの差分を抽出し、1 8 の下線部分の結果を得る。

【 0 0 2 0 】

要約 (重要文抽出) は、文書内の文字列から重要と思われる文を抽出する手法であり、例えば、特開平 1 1 - 2 7 2 6 8 6 号公報に開示されるものを適用できる。前処理部 3 では重要と思われる文字列 (文) を差分検出部 4 へ送る。

## 【 0 0 2 1 】

このときの処理イメージを図 6 に示す。19 と 20 はそれぞれ更新前と更新後の Web ページの要約抽出結果を示している。19 と 20 において、それぞれ重要でないと判断された文字列は二重線で消去している。差分検出部 4 では、重要文として抽出された 21 に示す文のみを比較することになり、下線で示す部位が差分であると抽出する。

## 【 0 0 2 2 】

前処理部 3 の別の手法としては、キーワード抽出による微小な修正等の除去も挙げることができる。キーワード抽出は、例えば、キーワードを「異なる文字コードに囲まれた漢字およびカナの連続文字」として定義することで、図 2 と図 3 の Web ページに対する処理結果は図 7 のようになる。フレーム 13、15 の変更箇所は、前記定義ではキーワードとはなり得ないため抽出されない。図 7 の内容を差分検出部 4 で比較することで、更新されたかどうかを判別することができる。仮に、フレーム 14 内の 1 月 1 日の記事における「します」を「しました」にする変更のみがあった場合、変更前と変更後のキーワードの違いは生じず、微小な修正であり、有意な更新ではないと判別される。

## 【 0 0 2 3 】

出力部 5 では、差分検出部 4 の結果に基づき、対象とする Web ページに有意な更新があることをユーザに通知する。

## 【 0 0 2 4 】

ユーザへの通知は、表示デバイスへの表示やメールによる通知などで行うことができ、通知内容には対象となった Web ページの URL や、変更を検知したフレームの情報でもよく、具体的な変更内容を含めてもかまわない。

## 【 0 0 2 5 】

なお、本説明では更新前の Web ページの情報をあらかじめ格納しておくバッファ、任意のタイミングで対象 Web ページを取得するタイマー類、および処理ステップ等は一般的であるため省略している。格納する Web ページの情報は、Web ページの生のデータでもかまわないし、前処理部 3 の処理後のデータでもかまわない。

【 0 0 2 6 】

( A - 3 ) 第 1 の実施形態の効果

以上の様に、第 1 の実施形態によれば、前処理部 3 において対象となる文書に対して主要個所の抽出を行うことで、差分検出部 4 では主要個所の変化を有意な更新として検知することが可能となる。

【 0 0 2 7 】

また、前処理部 3 においてキーワード抽出を行うことで、差分検出部 4 では微小な修正は検知対象外とすることにより、有意な更新のみを検知することが可能となる。

【 0 0 2 8 】

これにより、有意更新検知部 2 では対象となる文書の有意な更新があったことを検知し、出力部 5 においてその旨をユーザに伝えることが可能となる。

【 0 0 2 9 】

( B ) 第 2 の実施形態

( B - 1 ) 第 2 の実施形態の構成

図 8 は、第 2 の実施形態の電子文書有意更新通知装置の機能的構成を示すブロック図である。

【 0 0 3 0 】

入力部 1 は、インターネットやイントラネット等から W e b ページやテキスト等を取得して入力データとするものである。

【 0 0 3 1 】

有意更新検知部（第 2 の実施形態の電子文書有意更新検知装置） 6 は、W e b ページ等の有意な更新を検知するものであり、その内部では、差分抽出部 4 による文書の差分抽出と、価値判断部 7 による抽出した差分が有意なものであるかの判断とを行うものである。

【 0 0 3 2 】

出力部 5 では、W e b ページ等に有意な更新であった場合に、その旨を表示装置に表示したり、ユーザにメールで更新内容を通知したりする。

【 0 0 3 3 】

これらは、1つのシステム（パソコンやプロバイダーサーバ）上で構築してもかまわないし、ネットワークで結ばれたサーバ間で協調動作するように構築してもかまわない。

#### 【 0 0 3 4 】

#### （B－2）第2の実施形態の動作

第2の実施形態の詳細な処理についても、上述した図2および図3で示す架空のWebページを例として取り上げて説明する。

#### 【 0 0 3 5 】

有意更新検知部6は、差分抽出部4と価値判断部7を含んでおり、差分抽出部4では文書全体の差分の抽出を行い、その結果を価値判断部7で判断する。価値判断部7の処理は、例えば差分量の閾値処理や、形態素解析等の自然言語処理による属性判断等である。

#### 【 0 0 3 6 】

以下、それぞれの手法を用いた例を図9で説明する。

#### 【 0 0 3 7 】

差分量の閾値処理では、例えば個々の差分の文字列長がある閾値を超えた場合に更新されたと判断することができる。仮に差分の文字列は10文字以上で有効とした場合、図9の例では差分22に対して結果23を得ることができる。この場合、結果23では、閾値未満の文字列を削除（二重線）することで、まとまりのある文に価値があると判別している。

#### 【 0 0 3 8 】

形態素解析では、図9に示すように差分検出部4から渡された差分22をいくつかのパーツに分類する。各パーツの属性について、更新通知の価値を判別していく。その判別結果を23に示している。この場合、結果23では、文を構成しないパーツ（例えば助詞や単独の名詞等）を不要パーツとして判別（二重線で削除）することで、まとまりのある文に価値があると判別している。

#### 【 0 0 3 9 】

以上の手法により価値判断部7は価値があると判断された文字列を出力部5へわたす。

【 0 0 4 0 】

出力部 5 の処理は第 1 の実施形態と同様である。

【 0 0 4 1 】

なお、本説明では更新前の W e b ページの情報をあらかじめ格納しておくバッファ、任意のタイミングで対象 W e b ページを取得するタイマー類、および処理ステップ等は一般的であるため省略している。

【 0 0 4 2 】

( B - 3 ) 第 2 の実施形態の効果

以上の様に、第 2 の実施形態の説明によれば、価値判断部 7 において対象となる文書の差分文字列に対して価値判断を行うことで、文書の小さな修正等を更新情報から除外することができる。

【 0 0 4 3 】

これにより有意更新検知部 6 では対象となる文書の更新内容のうち有意な情報のみを検知し、出力部 5 においてその更新内容をユーザに伝えることが可能となる。

【 0 0 4 4 】

( C ) 他の実施形態

第 1 の実施形態および第 2 の実施形態は、インターネット、イントラネットの W e b ページやテキスト文書を監視するシステムに利用することができ、システム側では多くのユーザからの個別アクセスによるトラフィックを減らすことができるとともに、ユーザ側ではサイトの巡回を行う時間と労力を削減することが可能となる。

【 0 0 4 5 】

第 1 の実施形態では主に有意な更新があったかどうかを検知してその旨を通知し、第 2 の実施形態は主に更新内容について有意と判断した情報そのものを通知することを特徴とする。

【 0 0 4 6 】

第 1 の実施形態と第 2 の実施形態は、個別に実装してもかまわないが、同時に実装してもよい。また、第 1 の実施形態の前処理部 3 で用いた処理を、第 2 の実

施形態の価値判別部 7 にアレンジ（あるいはその逆）してもかまわない。これらの工夫で処理の強化や、サイト毎のきめ細かな処理に対応することが可能となる。

【 0 0 4 7 】

【発明の効果】

以上のように、本発明によれば、電子文書の更新と呼ぶに相応しいようなレベルの更新を検知したり通知したりできるようになる。

【図面の簡単な説明】

【図 1】

第 1 の実施形態の電子文書有意更新通知装置の機能的構成を示すブロック図である。

【図 2】

更新前の W e b ページ例を示す説明図である。

【図 3】

図 2 の W e b ページに対応した更新後の W e b ページ例を示す説明図である。

【図 4】

第 1 の実施形態のフレーム事前指定に利用する注目箇所テーブルを示す説明図である。

【図 5】

第 1 の実施形態の W e b ページでの注目フレームの説明図である。

【図 6】

第 1 の実施形態の重要文の抽出方法の説明図である。

【図 7】

第 1 の実施形態のキーワード抽出という前処理で適用するキーワード例の説明図である。

【図 8】

第 2 の実施形態の電子文書有意更新通知装置の機能的構成を示すブロック図である。

【図 9】



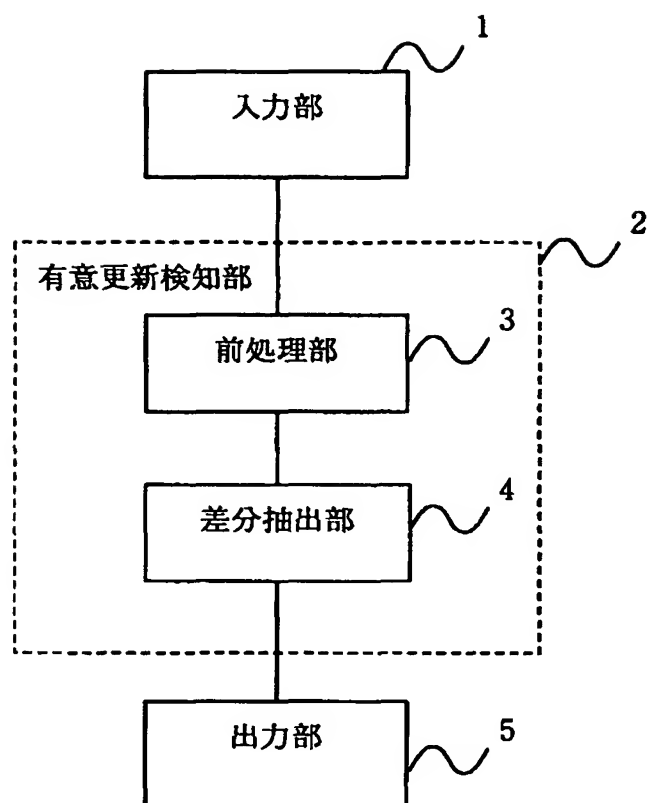
第 2 の実施形態の動作の説明図である。

【符号の説明】

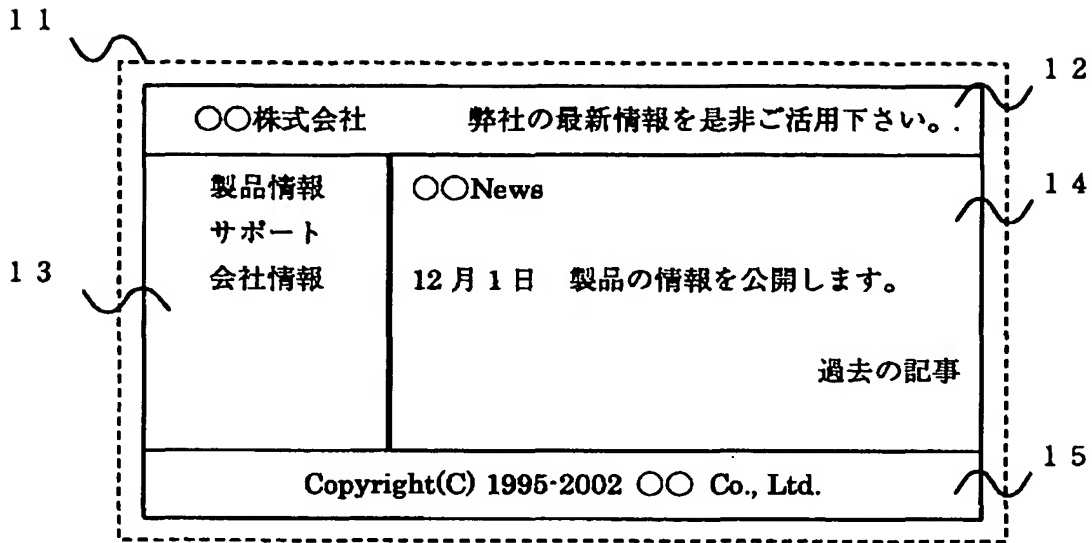
1 …入力部、2、6 …有意更新検知部（電子文書有意更新検知装置）、3 …前  
処理部、4 …差分抽出部、5 …出力部、7 …価値判断部。

【書類名】 図面

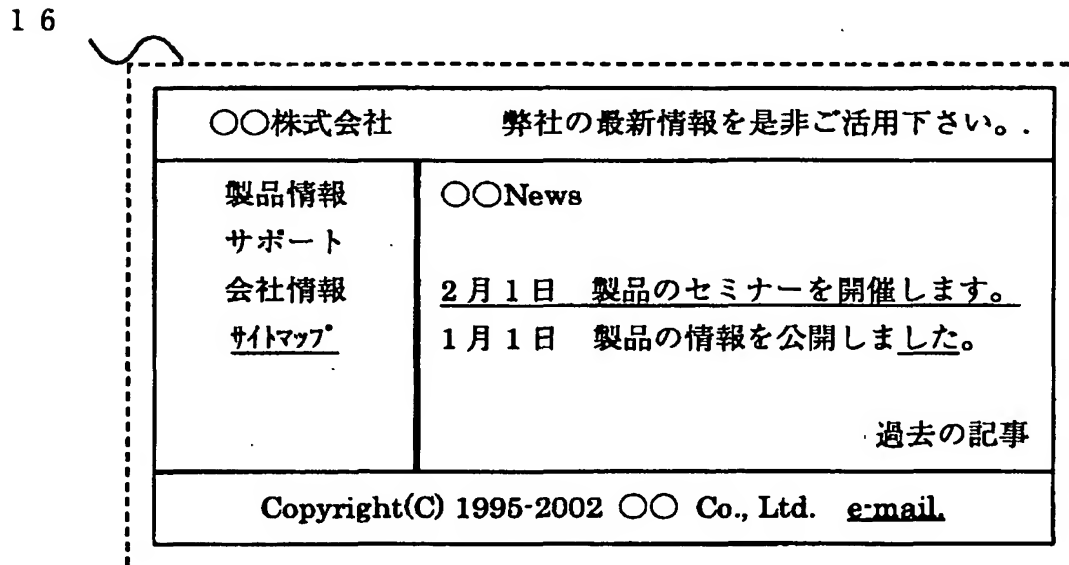
【図 1】



【図 2】



【図 3】



【図4】

URL	frame
http://www.〇〇.jp/	News.html
http://...	...

【図5】

〇〇株式会社
弊社の最新情報を是非ご利用下さい。

製品情報

サポート

会社情報

11月32'

〇〇News

2月1日 製品のセミナーを開催します。  
1月1日 製品の情報を公開しました。

過去の記事

Copyright(©) 1995-2002 〇〇 Co., Ltd. e-mail

17

18

【図 6】

〇〇株式会社 弊社の最新情報を是非ご活用下さい。.		19
製品情報	〇〇News	
<del>サポ</del>	1月1日 製品の情報を公開します。	
会社情報	過去の記事	
Copyright(C) 1995-2002 〇〇 Co., Ltd.		

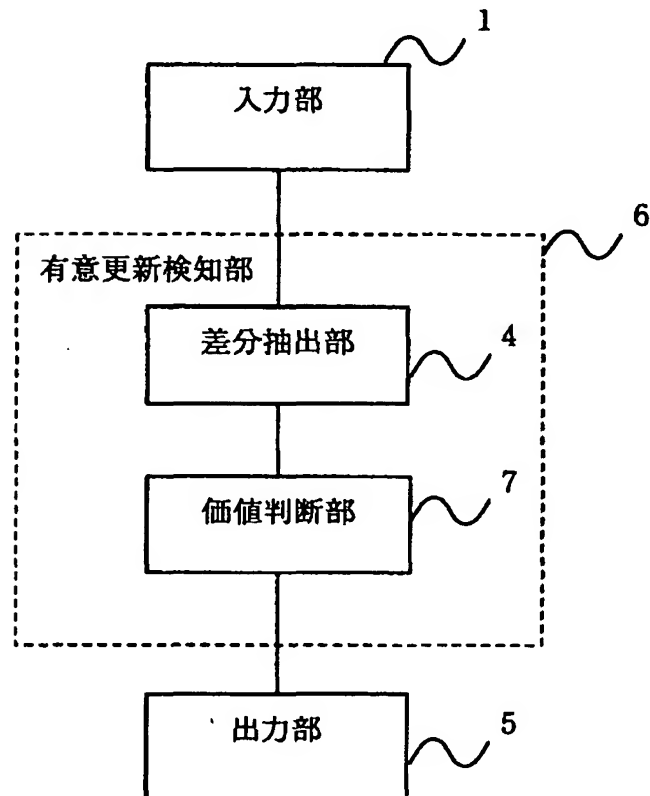
〇〇株式会社 弊社の最新情報を是非ご活用下さい。.		20
製品情報	〇〇News	
<del>サポ</del>	2月1日 製品のセミナーを開催します。	
会社情報	1月1日 製品の情報を公開しました。	
<del>お問い合わせ</del>	過去の記事	
Copyright(C) 1995-2002 〇〇 Co., Ltd. e-mail		

弊社の最新情報を是非ご活用下さい。.		21
2月1日 製品のセミナーを開催します。		
1月1日 製品の情報を公開しました。		

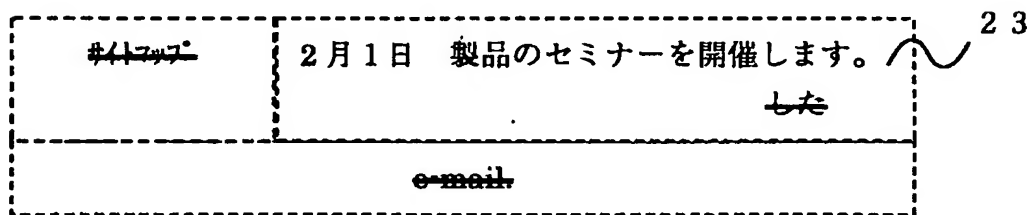
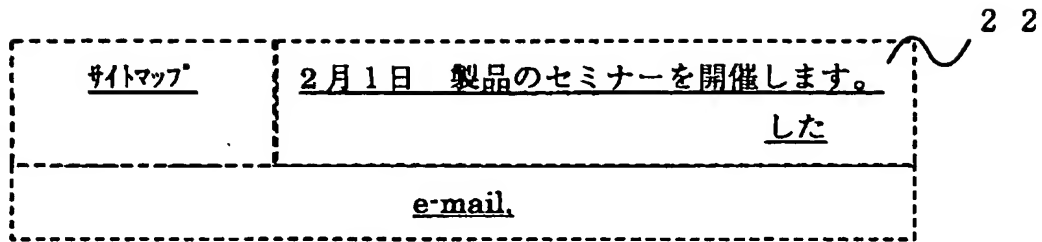
【図 7】

更新前のキーワード	更新後のキーワード
製品	製品
情報	情報
公開	公開
弊社	弊社
最新情報	最新情報
活用	活用
	セミナー
	開催

【図 8】



【図9】



【書類名】 要約書

【要約】

【課題】 電子文書の更新と呼ぶに相応しいようなレベルの更新を検知したり通知したりできる電子文書有意更新検知装置や電子文書有意更新通知装置を提供する。

【解決手段】 本発明の電子文書有意更新検知装置は、複数の電子文書のそれぞれに対して、重要な箇所を抽出する１つ以上の前処理手段と、前処理により抽出された結果に対して差分検知を行う差分検知手段とを備える。又は、複数の電子文書それぞれに対して差分検知を行う差分検知手段と、抽出された差分に対して有意かどうかの価値を判断する１つ以上の価値判断手段とを備える。又は、複数の電子文書のそれぞれに対して重要な箇所を抽出する１つ以上の前処理手段と、それぞれの電子文書に対して差分検知を行う差分検知手段と、抽出された差分に対して有意かどうかの価値を判断する１つ以上の価値判断手段とを備える。本発明の電子文書有意更新通知装置は、本発明の電子文書有意更新検知装置を備える。

【選択図】 図 1



出 願 人 履 歴 情 報

識別番号 [ 0 0 0 0 0 0 2 9 5 ]

1. 変更年月日	1 9 9 0 年 8 月 2 2 日
[変更理由]	新規登録
住 所	東京都港区虎ノ門1丁目7番12号
氏 名	沖電気工業株式会社